

УДК 519.713

О МАТРИЦЕ ПЕРВЫХ МОМЕНТОВ РАЗЛОЖИМОЙ СТОХАСТИЧЕСКОЙ КС-ГРАММАТИКИ

Л.П. Жильцова

Аннотация

Рассматривается стохастическая контекстно-свободная грамматика с произвольным числом классов нетерминальных символов без ограничений на порядок следования классов. Соответствующая ей матрица A первых моментов является разложимой. Для случая, когда перронов корень r матрицы A строго меньше единицы, исследуются свойства матрицы A^t при $t \rightarrow \infty$.

Ключевые слова: алгоритм, кодирование, сжатие, контекстно-свободный язык, грамматика, оптимизация, автомат, вероятность.

Введение

В работах [1, 2] нами рассматривались вопросы, связанные с кодированием сообщений, являющихся словами стохастического контекстно-свободного языка (стохастического КС-языка), при условии, что матрица первых моментов грамматики неразложима, непериодична и ее максимальный по модулю собственный корень (перронов корень) строго меньше единицы (докритический случай). При неразложимой матрице первых моментов нетерминальные символы грамматики образуют один класс.

При изучении вопросов кодирования важную роль играет строение матрицы A первых моментов и асимптотическое поведение элементов матрицы A^t при $t \rightarrow \infty$. Свойства неразложимой матрицы A^t изучены в [3]. Для докритического случая в [4] проведено исследование грамматик с двумя классами нетерминальных символов, при этом установлено асимптотическое поведение матрицы A^t .

В настоящей работе свойства матриц A и A^t исследуются для стохастических КС-грамматик с произвольным числом классов нетерминальных символов грамматики без ограничений на порядок следования классов.

1. Основные определения

Для изложения результатов о контекстно-свободных языках будем использовать определения КС-языка и стохастического КС-языка из [5, 6]. Стохастической КС-грамматикой называется система $G = \langle V_T, V_N, R, s \rangle$, где V_T и V_N – конечные множества терминальных и нетерминальных символов (терминалов и нетерминалов) соответственно; $s \in V_N$ – аксиома, R – множество правил. Множество R можно представить в виде $R = \bigcup_{i=1}^k R_i$, где k – мощность алфавита V_N и $R_i = \{r_{i1}, \dots, r_{in_i}\}$. Каждое правило r_{ij} из R_i имеет вид

$$r_{ij} : A_i \xrightarrow{p_{ij}} \beta_{ij}, \quad j = 1, \dots, n_i,$$

где $A_i \in V_N, \beta_{ij} \in (V_T \cup V_N)^*$ и p_{ij} – вероятность применения правила r_{ij} (вероятность правила r_{ij}), которая удовлетворяет следующим условиям:

$$0 < p_{ij} \leq 1 \quad \text{и} \quad \sum_{j=1}^{n_i} p_{ij} = 1, \quad i = 1, 2, \dots, k.$$

Для слов α и β из $(V_T \cup V_N)^*$ будем говорить, что β непосредственно выводимо из α (и записывать $\alpha \Rightarrow \beta$), если $\alpha = \alpha_1 A_i \alpha_2, \beta = \alpha_1 \beta_{ij} \alpha_2$ для некоторых $\alpha_1, \alpha_2 \in (V_T \cup V_N)^*$, и в грамматике G имеется правило $A_i \xrightarrow{p_{ij}} \beta_{ij}$.

Обозначим через \Rightarrow_* рефлексивное транзитивное замыкание отношения \Rightarrow . КС-язык, порождаемый грамматикой G , определяется как множество слов $L_G = \{\alpha : s \Rightarrow_* \alpha, \alpha \in V_T^*\}$.

Каждому слову α КС-языка соответствует последовательность правил грамматики (вывод), с помощью которой α выводится из аксиомы s . Вероятность вывода определяется как произведение вероятностей правил, образующих вывод. Вероятность слова α определяется как сумма вероятностей всех различных левых выводов слова α (при левом выводе очередное правило применяется к самому левому нетерминалу в слове).

Грамматика G называется согласованной, если

$$\lim_{N \rightarrow \infty} \sum_{\alpha \in L, |\alpha| \leq N} p(\alpha) = 1$$

(здесь $|x|$ – длина слова x). В работе рассматриваются согласованные КС-грамматики. Согласованная КС-грамматика G индуцирует распределение вероятностей P на множестве слов порождаемого КС-языка L и определяет стохастический КС-язык $\mathcal{L} = (L, P)$.

Важное значение имеет понятие дерева вывода. Дерево строится следующим образом.

Корень дерева помечается аксиомой s . Пусть при выводе слова α на очередном шаге в процессе левого вывода применяется правило $A \xrightarrow{p_{ij}} b_{i_1} b_{i_2} \dots b_{i_m}$, где $b_{i_l} \in V_N \cup V_T$ ($l = 1, \dots, m$). Тогда из самой левой вершины-листа дерева, помеченной символом A (при обходе листьев дерева слева направо), проводится m дуг в вершины следующего яруса, которые помечаются слева направо символами $b_{i_1}, b_{i_2}, \dots, b_{i_m}$ соответственно. После построения дуг и вершин для всех правил грамматики в выводе слова языка все листья дерева помечены терминальными символами и само слово получается при обходе листьев дерева слева направо.

Ярусы дерева будем нумеровать следующим образом. Корень дерева расположен в нулевом ярусе. Вершины дерева, смежные с корнем, образуют первый ярус, и т. д. Дуги, выходящие из вершин j -го яруса, ведут к вершинам $(j+1)$ -го яруса.

Рассмотрим многомерные производящие функции $F_i(s_1, s_2, \dots, s_k)$, $i = 1, \dots, k$, где переменная s_i соответствует нетерминальному символу A_i . Функция $F_i(s_1, s_2, \dots, s_k)$ строится по множеству правил R_i с одинаковой левой частью A_i следующим образом.

Для каждого правила $A_i \xrightarrow{p_{ij}} \beta_{ij}$ выписывается слагаемое

$$q_{ij} = p_{ij} \cdot s_1^{l_1} \cdot s_2^{l_2} \cdot \dots \cdot s_k^{l_k},$$

где l_m – число вхождений нетерминального символа A_m в правую часть правила ($m = 1, \dots, k$). Тогда

$$F_i(s_1, s_2, \dots, s_k) = \sum_{j=1}^{n_i} q_{ij}.$$

Пусть

$$a_j^i = \left. \frac{\partial F_i(s_1, \dots, s_k)}{\partial s_j} \right|_{s_1=s_2=\dots=s_k=1}.$$

Квадратная матрица A порядка k , образованная элементами a_j^i , называется матрицей первых моментов грамматики G .

Так как матрица A неотрицательна, существует максимальный по модулю действительный неотрицательный собственный корень (перронов корень) [7]. Обозначим этот корень через r .

Известно необходимое и достаточное условие согласованности стохастической КС-грамматики [6]: стохастическая КС-грамматика при отсутствии бесполезных нетерминалов (то есть не участвующих в порождении слов языка) является согласованной тогда и только тогда, когда перронов корень матрицы первых моментов не превосходит единицы.

В работе рассматривается случай, когда $r < 1$.

Введем некоторые обозначения. Будем говорить, что нетерминал A_j непосредственно следует за нетерминалом A_i (и обозначать $A_i \rightarrow A_j$), если в грамматике существует правило вида $A_i \xrightarrow{p_{ij}} \alpha_1 A_j \alpha_2$, где $\alpha_1, \alpha_2 \in (V_T \cup V_N)^*$. Транзитивное замыкание отношения \rightarrow обозначим \rightarrow_* .

Пусть $A_i \in V_N$. Через $I_1(A_i)$ обозначим множество нетерминалов таких, что $A_i \rightarrow_* A_j$ для любого $A_j \in I_1(A_i)$. Через $I_2(A_i)$ обозначим множество нетерминалов таких, что $A_j \rightarrow A_i$ для любого $A_j \in I_2(A_i)$. Через $I_0(A_i)$ обозначим пересечение этих множеств, то есть $I_0(A_i) = I_1(A_i) \cap I_2(A_i)$. Множество нетерминалов $K = \{A_{i_1}, \dots, A_{i_q}\}$, для которых $I_0(A_{i_j})$ совпадают и $I_0(A_{i_j}) \neq \emptyset$, $j = 1, \dots, q$, назовем классом. Если $I_0(A_i) = \emptyset$, то будем считать, что A_i образует особый класс $\{A_i\}$.

Грамматика G называется неразложимой, если все нетерминалы из V_N образуют один класс. В противном случае G называется разложимой. Разложимой грамматике соответствует разложимая матрица [7] первых моментов.

Для различных классов нетерминалов K_1 и K_2 будем говорить, что класс K_2 непосредственно следует за классом K_1 (и обозначать $K_1 \prec K_2$), если существуют $A_1 \in K_1$ и $A_2 \in K_2$ такие, что $A_1 \rightarrow A_2$. Рефлексивное транзитивное замыкание отношения \prec обозначим через \prec_* и назовем отношением следования.

Пусть $\{K_1, K_2, \dots, K_m\}$ – множество классов нетерминалов грамматики, $m \geq 2$. Без ограничения общности можно считать, что в грамматике нет особых классов.

Будем полагать, что классы нетерминалов перенумерованы таким образом, что $K_i \prec K_j$ тогда и только тогда, когда $i < j$ (это всегда можно сделать). Соответствующая разложимой грамматике матрица первых моментов A имеет следующий вид:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1m-1} & A_{1m} \\ 0 & A_{22} & \dots & A_{2m-1} & A_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{m-1m-1} & A_{m-1m} \\ 0 & 0 & \dots & 0 & A_{mm} \end{pmatrix}. \quad (1)$$

Один класс нетерминалов в матрице первых моментов представлен множеством подряд идущих строк и соответствующим множеством столбцов с теми же номерами. Для класса K_i квадратная подматрица, образованная соответствующими строками и столбцами, обозначается через A_{ii} . Подматрица A_{ij} является нулевой, если $K_i \not\prec K_j$. Блоки, расположенные ниже главной диагонали, – нулевые в силу упорядоченности классов нетерминалов.

Для каждого класса K_i матрица A_{ii} неразложима. Без ограничения общности будем считать, что она строго положительна и неперiodична. Этого всегда можно добиться, применяя метод укрупнения правил грамматики, описанный в [2].

Обозначим через r_i перронов корень матрицы A_{ii} . Для неразложимой матрицы перронов корень является действительным и простым [7]. Очевидно, в силу структуры матрицы первых моментов, $r = \max_i \{r_i\}$ и $r > 0$.

Пусть $J = \{i_1, i_2, \dots, i_l\}$ – множество всех номеров i_j классов, для которых $r_{i_j} = r$.

Введем некоторые соглашения. Рассматривая вектор, будем считать, что мы имеем дело с вектором-столбцом, если противное не оговорено специально. В дальнейшем для вектора или матрицы X будем писать $X = c$ ($X \leq c$), где c – скаляр, если все компоненты вектора или матрицы X равны (соответственно меньше или равны) c . Через X^T обозначим транспонирование вектора или матрицы X . Через (v_1, v_2) , где v_1, v_2 – векторы-столбцы, будем обозначать объединенный вектор-столбец $(v_1^T, v_2^T)^T$.

2. Асимптотика для матрицы A^t

Зафиксируем пару (l, h) , $l, h \in \{1, 2, \dots, m\}$, и рассмотрим всевозможные последовательности классов $K_{i_1} \prec K_{i_2} \prec \dots \prec K_{i_s}$, где $i_1 = l, i_s = h$. Среди всех таких последовательностей выберем ту, которая содержит наибольшее число классов с номерами из J . Это число обозначим через s_{lh} .

Дополнительно переупорядочим классы по неубыванию величины s_{1l} , причем при одинаковых значениях s_{1l} сначала поставим классы с номерами из J .

Разобьем последовательность классов на группы классов $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_w$, при этом класс K_l отнесем к группе \mathcal{M}_1 при $s_{1l} \leq 1$ и к группе \mathcal{M}_j при $s_{1l} = j$ ($l = 2, \dots, m$). Для краткости далее будем называть группу классов просто группой.

Для групп \mathcal{M}_i и \mathcal{M}_j определим s_{ij}^* как $\max_{K_l \in \mathcal{M}_i, K_h \in \mathcal{M}_j} \{s_{lh}\}$.

Матрицу первых моментов будем также представлять в виде

$$A = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1w} \\ 0 & B_{22} & \dots & B_{2w} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & B_{ww} \end{pmatrix},$$

где B_{ij} – подматрица на пересечении строк для классов из группы \mathcal{M}_i и столбцов для классов из \mathcal{M}_j . Очевидно, каждая матрица B_{ii} имеет перронов корень r .

Рассмотрим подматрицу B_{lh} матрицы первых моментов A . Запись $B_{lh}^{(t)}$ будем применять для обозначения соответствующей подматрицы матрицы A^t . Изучим поведение матрицы A^t при $t \rightarrow \infty$.

Теорема 1. При $t \rightarrow \infty$

$$B_{lh}^{(t)} = H_{lh} \cdot t^{s_{lh}^* - 1} r^t (1 + o(1)),$$

где H_{lh} – матрица, не зависящая от t .

Доказательство. Доказательство проведем индукцией по числу групп.

Пусть $A = (B_{11})$. Группу \mathcal{M}_1 разобьем на три подгруппы. К первой подгруппе \mathcal{M}_{11} отнесем классы нетерминалов с $s_{1l} = 0$, ко второй подгруппе \mathcal{M}_{12} – классы с номерами из множества J , к третьей подгруппе \mathcal{M}_{13} – все остальные классы.

В соответствии с этим разбиением B_{11} представим в виде

$$B_{11} = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix},$$

где C_{ij} – подматрица со строками для классов из \mathcal{M}_{1i} и столбцами для классов из \mathcal{M}_{1j} .

Исследуем строение подматриц матрицы B_{11}^t при $t \rightarrow \infty$. Заметим, что подгруппа \mathcal{M}_{11} может быть пустой, тогда в B_{11} присутствуют только подматрицы C_{22} , C_{23} и C_{33} .

Известно следующее представление для степени произвольной матрицы C [7]:

$$C^t = \sum_{l=1}^s \left(\lambda_l^t Z_{l1} + (\lambda_l^t)' \cdot Z_{l2} + \dots + (\lambda_l^t)^{(m_l-1)} Z_{lm_l} \right), \quad (2)$$

где λ_l – корни минимального многочлена $\psi(\lambda)$ матрицы C ($l = 1, \dots, s$), m_l – кратность корня λ_l для минимального многочлена, $(\lambda_l^t)^{(n)}$ – n -я производная по λ_l от λ_l^t , матрицы Z_{lj} вполне определяются заданием матрицы C и не зависят от t .

Так как каждому классу K_i из \mathcal{M}_{11} или из \mathcal{M}_{13} соответствует перронов корень $r_i < r$, то для C_{11}^t и C_{33}^t из (2) следуют оценки $C_{11}^t = o(r^t)$ и $C_{33}^t = o(r^t)$.

Пусть \mathcal{M}_{1l} содержит j_l классов ($l = 1, 2, 3$). Любому классу K_i из \mathcal{M}_{12} соответствует неразложимая подматрица A_{ii} в представлении (1), и классы из \mathcal{M}_{12} попарно несравнимы.

Для неразложимой положительной матрицы A_{ii}^t применим представление, установленное в [3]:

$$A_{ii}^t = U_i V_i \cdot r^t (1 + o(1)),$$

где U_i – правый собственный вектор (вектор-столбец), V_i – левый собственный вектор (вектор-строка), соответствующие r , $U_i > 0$, $V_i > 0$ и $V_i U_i = 1$.

Поэтому матрицу C_{22}^t можно представить в следующем виде:

$$C_{22}^t = \begin{pmatrix} U_{j_1+1} V_{j_1+1} & 0 & \dots & 0 & 0 \\ 0 & U_{j_1+2} V_{j_1+2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & U_{j_1+j_2} V_{j_1+j_2} \end{pmatrix} \cdot r^t (1 + o(1)). \quad (3)$$

Обозначим через D матрицу

$$\begin{pmatrix} U_{j_1+1} V_{j_1+1} & 0 & \dots & 0 & 0 \\ 0 & U_{j_1+2} V_{j_1+2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & U_{j_1+j_2} V_{j_1+j_2} \end{pmatrix}.$$

Очевидно, что D можно представить в виде $\sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(2)}$, где $U_i^{(2)} = (0, \dots, 0, U_i, 0, \dots, 0)$, $V_i^{(2)} = (0, \dots, 0, V_i, 0, \dots, 0)$ и U_i и V_i расположены на местах, соответствующих классу K_i .

Непосредственной проверкой устанавливается, что

$$C_{12}^{(t)} = \sum_{j=0}^{t-1} C_{11}^j C_{12} C_{22}^{t-j-1} = \sum_{j=0}^{t-\lfloor \log \log t \rfloor} C_{11}^j C_{12} C_{22}^{t-j-1} + \sum_{j=t-\lfloor \log \log t \rfloor+1}^{t-1} C_{11}^j C_{12} C_{22}^{t-j-1}.$$

Обозначим $\sum_{j=0}^{t-\lfloor \log \log t \rfloor} C_{11}^j C_{12} C_{22}^{t-j-1}$ через Σ_1 и $\sum_{j=t-\lfloor \log \log t \rfloor+1}^{t-1} C_{11}^j C_{12} C_{22}^{t-j-1}$ через Σ_2 .

Оценим сначала Σ_2 . Для этого применим представление (2) для C_{11}^t . Отметим, что для C_{11} все собственные корни строго меньше r , а элементы матрицы C_{22}^t ограничены некоторой константой, не зависящей от t . Поэтому

$$\Sigma_2 \leq c_1 \cdot (r')^{t - \lfloor \log \log t \rfloor + 1} \sum_{j=t - \lfloor \log \log t \rfloor + 1}^{t-1} C_{12} C_{22}^{t-j-1} \leq$$

$$c_2 \cdot (r')^{t - \lfloor \log \log t \rfloor + 1} \cdot \lfloor \log \log t \rfloor = O(\log \log t \cdot (\log t)^{c_3} \cdot (r')^t) = o(r^t)$$

для некоторых констант c_1 , c_2 и c_3 , и r' , удовлетворяющего неравенству $r' < r$.

Учитывая представление (3) для C_{22}^{t-j-1} , получим:

$$\Sigma_1 = r^{t-1} (1 + o(1)) \sum_{j=0}^{t - \lfloor \log \log t \rfloor} \left(\frac{C_{11}}{r} \right)^j C_{12} D.$$

Нетрудно проверить, что

$$\frac{1}{r} \sum_{j=0}^{\infty} \left(\frac{C_{11}}{r} \right)^j = (rE - C_{11})^{-1},$$

поскольку перронов корень матрицы C_{11} строго меньше r , и, следовательно, обратная матрица для $(rE - C_{11})$ существует. Поэтому

$$\Sigma_1 = r^t (rE - C_{11})^{-1} C_{12} D \cdot (1 + o(1)).$$

Найдем правый собственный вектор матрицы B_{11} для перронова корня r . Компоненты вектора представим в виде $U = (U^{(1)}, U^{(2)}, U^{(3)})$, где $U^{(1)}$ соответствует \mathcal{M}_{11} , $U^{(2)} - \mathcal{M}_{12}$, и $U^{(3)} - \mathcal{M}_{13}$. Вектор U удовлетворяет системе уравнений

$$\begin{cases} C_{11}U^{(1)} + C_{12}U^{(2)} + C_{13}U^{(3)} = rU^{(1)}, \\ C_{22}U^{(2)} + C_{23}U^{(3)} = rU^{(2)}, \\ C_{33}U^{(3)} = rU^{(3)}. \end{cases}$$

Отсюда $U^{(3)} = 0$, так как перронов корень матрицы C_{33} строго меньше r , $U^{(2)}$ является правым собственным вектором для матрицы C_{22} и $U^{(1)} = (rE - C_{11})^{-1} C_{12} U^{(2)}$.

Таким образом,

$$C_{12}^{(t)} = r^t (rE - C_{11})^{-1} C_{12} \sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(2)} \cdot (1 + o(1)) = r^t \sum_{i=j_1+1}^{j_2} U_i^{(1)} V_i^{(2)} \cdot (1 + o(1)),$$

где $U_i^{(1)}$ соответствует правому собственному вектору $U_i^{(2)}$ матрицы C_{22} .

Рассмотрим матрицу $C_{23}^{(t)}$. Непосредственной проверкой устанавливается, что

$$C_{23}^{(t)} = \sum_{j=0}^{t-1} C_{22}^j C_{23} C_{33}^{t-j-1}.$$

Представим эту сумму в виде $\Sigma_1 + \Sigma_2$, где

$$\Sigma_1 = \sum_{j=\lfloor \log \log t \rfloor + 1}^{t-1} C_{22}^j C_{23} C_{33}^{t-j-1}, \quad \Sigma_2 = \sum_{j=0}^{\lfloor \log \log t \rfloor} C_{22}^j C_{23} C_{33}^{t-j-1}.$$

Аналогично тому, как это сделано для C_{12}^t , доказывается оценка $\Sigma_2 = o(r^t)$. Для Σ_1 справедливы соотношения

$$\Sigma_1 = D \cdot \sum_{j=\lfloor \log \log t \rfloor + 1}^{t-1} r^j C_{23} C_{33}^{t-j-1} \cdot (1 + o(1)) = r^t \cdot (1 + o(1)) \sum_{j=0}^{t-\lfloor \log \log t \rfloor - 2} \left(\frac{C_{33}}{r} \right)^j.$$

Найдем левый собственный вектор матрицы B_{11} для перронова корня r . Компоненты вектора представим в виде $V = (V^{(1)}, V^{(2)}, V^{(3)})$. Вектор V удовлетворяет системе уравнений

$$\begin{cases} V^{(1)} C_{11} = r V^{(1)}, \\ V^{(1)} C_{12} + V^{(2)} C_{22} = r V^{(2)}, \\ V^{(1)} C_{13} + V^{(2)} C_{23} + V^{(3)} C_{33} = r V^{(3)}. \end{cases}$$

Отсюда $V^{(1)} = 0$, так как перронов корень матрицы C_{11} строго меньше r , $V^{(2)}$ является левым собственным вектором для матрицы C_{22} , и $V^{(3)} = (V^{(2)} C_{23} (rE - C_{33})^{-1})$. Поэтому

$$C_{23}^{(t)} = r^t \sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(2)} C_{23} (rE - C_{33})^{-1} \cdot (1 + o(1)) = r^t \sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(3)} \cdot (1 + o(1)).$$

Наконец, рассмотрим матрицу $C_{13}^{(t)}$. Нетрудно проверить, что

$$C_{13}^{(t)} = \sum_{j=0}^{t-1} C_{11}^j C_{13} C_{33}^{t-j-1} + \sum_{j=0}^{t-1} C_{12}^j C_{23} C_{33}^{t-j-1}.$$

Так как перроновы корни матриц C_{11} и C_{33} строго меньше r , то справедлива оценка

$$\sum_{j=0}^{t-1} C_{11}^j C_{13} C_{33}^{t-j-1} = o(r^t).$$

Поэтому

$$C_{13}^{(t)} = \sum_{j=0}^{t-1} C_{12}^{(j)} C_{23} C_{33}^{t-j-1} + o(r^t).$$

Применяя полученную оценку для $C_{12}^{(j)}$, после несложных преобразований получим, что

$$C_{13}^{(t)} = r^t \cdot \sum_{i=j_1+1}^{j_2} U_i^{(1)} V_i^{(2)} C_{23} (rE - C_{33})^{-1} \cdot (1 + o(1)) = r^t \sum_{i=j_1+1}^{j_2} U_i^{(1)} V_i^{(3)} \cdot (1 + o(1)).$$

Таким образом,

$$B_{11}^t = \begin{pmatrix} 0 & \sum_{i=j_1+1}^{j_2} U_i^{(1)} V_i^{(2)} & \sum_{i=j_1+1}^{j_2} U_i^{(1)} V_i^{(3)} \\ 0 & \sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(2)} & \sum_{i=j_1+1}^{j_2} U_i^{(2)} V_i^{(3)} \\ 0 & 0 & 0 \end{pmatrix} \cdot r^t + o(r^t). \quad (4)$$

Отметим, что строки матрицы B_{11} , соответствующие классам $K_i \in \mathcal{M}_{12}$, то есть классам, для которых $i \in J$, пропорциональны компонентам правого собственного вектора $U_i^{(2)}$, а столбцы, соответствующие классам $K_j \in \mathcal{M}_{12}$, пропорциональны компонентам левого собственного вектора $V_j^{(2)}$.

Рассмотрим случай $w = 2$. В этом случае матрица A^t имеет вид

$$A^t = \begin{pmatrix} B_{11}^t & B_{12}^{(t)} \\ 0 & B_{22}^t \end{pmatrix}.$$

Строение матрицы B_{22}^t аналогично строению исследованной ранее подматрицы для B_{11}^t , соответствующей подклассам \mathcal{M}_{12} и \mathcal{M}_{13} . Для $B_{12}^{(t)}$ справедлива формула

$$B_{12}^{(t)} = \sum_{j=0}^{t-1} B_{11}^j B_{12} B_{22}^{t-j-1}.$$

Для того чтобы различать собственные вектора матриц B_{11}^t и B_{22}^t , введем второй верхний индекс, равный 1, в случае матрицы B_{11}^t , и равный 2, в случае матрицы B_{22}^t . Для группы \mathcal{M}_2 отсутствует подгруппа \mathcal{M}_{21} , поэтому $\mathcal{M}_2 = (\mathcal{M}_{22}, \mathcal{M}_{23})$. В результате преобразований, учитывающих вид (4), получим, что

$$B_{12}^{(t)} = \begin{pmatrix} 0 & \sum_i U_i^{(11)} V_i^{(21)} & \sum_i U_i^{(11)} V_i^{(31)} \\ 0 & \sum_i U_i^{(21)} V_i^{(21)} & \sum_i U_i^{(21)} V_i^{(31)} \\ 0 & 0 & 0 \end{pmatrix} \cdot B_{12} \times \\ \times \begin{pmatrix} \sum_j U_j^{(22)} V_j^{(22)} & \sum_j U_j^{(22)} V_j^{(32)} \\ 0 & 0 \end{pmatrix} \cdot tr^t + o(tr^t).$$

Представим матрицу B_{12} в виде

$$B_{12} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \\ D_{31} & D_{32} \end{pmatrix}, \quad (5)$$

где разбиение по строкам сделано в соответствии с подгруппами группы \mathcal{M}_1 , а по столбцам – в соответствии с подгруппами группы \mathcal{M}_2 . Тогда

$$B_{12}^{(t)} = \begin{pmatrix} \sum_{i,j} c_{ij} U_i^{(11)} V_j^{(22)} & \sum_{i,j} c_{ij} U_i^{(11)} V_j^{(32)} \\ \sum_{i,j} c_{ij} U_i^{(21)} V_j^{(22)} & \sum_{i,j} c_{ij} U_i^{(21)} V_j^{(32)} \\ 0 & 0 \end{pmatrix} \cdot tr^t + o(tr^t),$$

где $c_{ij} = (V_i^{(21)} D_{21} + V_i^{(31)} D_{31}) U_j^{(22)}$. Коэффициент $c_{ij} > 0$ в случае, когда $K_i \prec K_j$.

Блоки матрицы $B_{12}^{(t)}$ можно также представить в виде, аналогичном (4), сменив порядок суммирования и проведя суммирование по i :

$$B_{12}^{(t)} = \begin{pmatrix} \sum_j U_j^{(11)} V_j^{(22)} & \sum_j U_j^{(11)} V_j^{(32)} \\ \sum_j U_j^{(21)} V_j^{(22)} & \sum_j U_j^{(21)} V_j^{(32)} \\ 0 & 0 \end{pmatrix} \cdot tr^t + o(tr^t),$$

где $U_j^{(l1)} = \sum_i c_{ij} U_i^{(l1)}$, $l = 1, 2$.

В силу строения собственных векторов, соответствующих подгруппам \mathcal{M}_{12} и \mathcal{M}_{22} , строки для класса $K_i \in \mathcal{M}_{12}$ пропорциональны компонентам правого собственного вектора матрицы A_{ii} , а столбцы для $K_j \in \mathcal{M}_{22}$ – компонентам левого собственного вектора матрицы A_{jj} .

Таким образом, доказана справедливость теоремы для $w = 2$.

Заметим, что $D_{32}^{(t)} = O(r^t)$ и $D_{33}^{(t)} = O(r^t)$, для получения этих оценок достаточно в качестве матрицы первых моментов рассмотреть матрицу без строк и столбцов, соответствующих подгруппам \mathcal{M}_{11} и \mathcal{M}_{12} . При этом D_{32} перейдет в подматрицу C_{12} , а D_{33} – в подматрицу C_{13} .

Предположим, что утверждение теоремы справедливо для $w - 1$ групп. Докажем его для w групп.

Обозначим через D_1 подматрицу матрицы A , соответствующую первым $w - 1$ группам, и через D_2 – подматрицу, соответствующую группам с номерами $2, 3, \dots, w$. Тогда матрицу первых моментов A можно представить в виде

$$A = \begin{pmatrix} D_1 & E_1 \\ 0 & B_{ww} \end{pmatrix}, \quad \text{где } E_1 = \begin{pmatrix} B_{1w} \\ \dots \\ B_{w-1,w} \end{pmatrix},$$

и в виде

$$A = \begin{pmatrix} B_{11} & E_2 \\ 0 & D_2 \end{pmatrix}, \quad \text{где } E_2 = (B_{12} \quad \dots \quad B_{1w}).$$

Очевидно, что

$$A^t = \begin{pmatrix} D_1^t & E_1^{(t)} \\ 0 & B_{ww}^t \end{pmatrix} = \begin{pmatrix} B_{11}^t & E_2^{(t)} \\ 0 & D_2^t \end{pmatrix},$$

где $E_1^{(t)}$, $E_2^{(t)}$ – подматрицы в A^t , соответствующие подматрицам E_1 и E_2 в A . Для матриц D_1 и D_2 утверждение теоремы справедливо по предположению индукции. Поэтому достаточно доказать теорему для подматрицы $B_{1w}^{(t)}$.

Непосредственной проверкой устанавливается, что

$$B_{1w}^{(t)} = \sum_{l=1}^{w-1} \sum_{j=0}^{t-1} B_{1l}^{(j)} B_{lw} B_{ww}^{t-j-1}.$$

По предположению индукции $B_{1l}^{(j)} = O(t^{s_{1l}^* - 1} r^t)$. Так как группы упорядочены по возрастанию s_{1l}^* , то определяющим в сумме является слагаемое $B_{1,w-1}^{(j)}$. Поэтому

$$B_{1w}^{(t)} = \sum_{j=0}^{t-1} B_{1w-1}^{(j)} B_{w-1,w} B_{ww}^{t-j-1} \cdot (1 + o(1)).$$

Раскрывая $B_{1w-1}^{(j)}$, после несложных преобразований получим:

$$B_{1w}^{(t)} = \begin{pmatrix} \sum_j U_j'^{(11)} V_j^{(2w)} & \sum_j U_j'^{(11)} V_j^{(3w)} \\ \sum_j U_j'^{(21)} V_j^{(2w)} & \sum_j U_j'^{(21)} V_j^{(3w)} \\ 0 & 0 \end{pmatrix} \cdot t^{s_{1w}^* - 1} r^t + o(t^{s_{1w}^* - 1} r^t),$$

где

$$U_j'^{(l1)} = \frac{1}{s_{1w}^* - 1} \sum_i (V_i^{(21)} D_{21} + V_i^{(31)} D_{31}) U_j^{(2w)} U_i^{(l1)}$$

и D_{l1} – блоки матрицы $B_{w-1,w}$ ($l = 1, 2$).

Теорема доказана. \square

Из доказательства теоремы вытекает

Следствие 1. При $t \rightarrow \infty$

$$B_{lh}^{(t)} = H_{lh} \cdot t^{s_{lh}^* - 1} r^t (1 + o(1)) \quad \text{для } l \neq h,$$

где

$$H_{lh} = \begin{pmatrix} \sum_j U_j^{(1l)} V_j^{(2h)} & \sum_j U_j^{(1l)} V_j^{(3h)} \\ \sum_j U_j^{(2l)} V_j^{(2h)} & \sum_j U_j^{(2l)} V_j^{(3h)} \\ 0 & 0 \end{pmatrix}.$$

Проинтерпретируем результаты теоремы 1, используя деревья вывода. Обозначим через D множество деревьев вывода, корень которых помечен аксиомой грамматики A_1 , и через $M_i(t)$ – математическое ожидание числа вершин на ярусе t дерева вывода из D , помеченных нетерминальным символом A_i .

Следствие 2. Пусть $A_i \in K_j$. Тогда при $t \rightarrow \infty$

$$M_i(t) \sim c_i \cdot t^{s_{1j}-1} r^t,$$

где c_i – некоторая неотрицательная константа.

Работа выполнена при финансовой поддержке РФФИ (проект № 07-01-00739).

Summary

L.P. Zhiltsova. On a Matrix of First Moments for Decomposable Stochastic CF-Grammar.

Stochastic context-free grammar is considered which contains arbitrary number of classes of non-terminal symbols without restrictions on the succession order of classes. Corresponding matrix A of first moments is decomposable. For the case when Perron's root of the matrix A is strictly less than one, properties of the matrix A^t are investigated under $t \rightarrow \infty$.

Key words: algorithm, coding, compression, context-free language, grammar, optimization, automaton, probability.

Литература

1. Жильцова Л.П. Закономерности применения правил грамматики в выводах слов стохастического контекстно-свободного языка // Матем. вопр. кибернетики. – 2000. – Вып. 9. – С. 101–126.
2. Жильцова Л.П. О нижней оценке стоимости кодирования и асимптотически оптимальном кодировании стохастического контекстно-свободного языка // Дискр. анализ и исслед. операций. Сер. 1. – 2001. – Т. 8, № 3. – С. 26–45.
3. Севастьянов Б.А. Ветвящиеся процессы. – М.: Наука, 1971. – 436 с.
4. Борисов А.Е. О свойствах стохастического КС-языка, порожденного грамматикой с двумя классами нетерминальных символов // Дискр. анализ и исслед. операций. Сер. 1. – 2005. – Т. 12, № 3. – С. 3–31.
5. Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. Т. 1. – М.: Мир, 1978. – 616 с.
6. Фу К. Структурные методы в распознавании образов. – М.: Мир, 1977. – 320 с.
7. Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1966. – 576 с.

Поступила в редакцию
30.03.09

Жильцова Лариса Павловна – доктор физико-математических наук, профессор кафедры математической логики и высшей алгебры Нижегородского государственного университета им. Н.И. Лобачевского.

E-mail: larzhil@rambler.ru